SafeTab: DP Algorithms for 2020 Census Detailed DHC Race & Ethnicity

Tumult Labs

March 9, 2022
v1.0.0

Abstract

This article describes the proposed differentially private (DP) algorithms that the US Census Bureau will use to release the Detailed Demographic and Housing Characteristics (DHC) Race & Ethnicity tabulations as part of the 2020 Decennial Census. The tabulations contain statistics (counts) of demographic and housing characteristics of the entire population of the US crossed with detailed races and tribes at varying levels of geography. We describe two differentially private algorithmic strategies, one based on adding noise drawn from a Geometric distribution that satisfies "pure"-DP, and another based on addition noise from a Discrete Gaussian distribution that satisfied a well studied variant of differential privacy, called Zero Concentrated Differential Privacy (zCDP). We analytically estimate the privacy loss parameters ensured by the two algorithms for comparable levels of error introduced in the statistics.

Contents

1 Executive Summary

In this article we describe SafeTab, a differentially private algorithm for releasing statistics about persons and households in the US that make up the Detailed DHC (Race & Ethnicity) data product (DHC-RE) to be released as part of the 2020 Decennial Census by the US Census Bureau.

The paper is organized as follows. Section 2 defines the DHC-RE data product and the private release problem that SafeTab solves. Section 3 describes the SafeTab algorithm. SafeTab can be instatiated in two ways:

- SafeTab[Geometric] that adds noise from the two sided Geometric distribution to the statistics that are released.
- SafeTab[Discrete Gaussian] that adds noise from a Discrete Gaussian distribution [5] to the statistics that are released.

Section 4 provides the necessary background on differential privacy and related definitions. We derive a couple of novel privacy results in Section 5 which are then used to prove the privacy properties of SafeTab[Geometric] (Section 6) and SafeTab[Discrete Gaussian] (Section 7). We provide a pure-DP and Renyi-DP analysis of SafeTab[Geometric], and a zCDP analysis of SafeTab[Discrete Gaussian].

In Section 8, we comapre the two SafeTab algorithms. Specifically, given accuracy requires of the output statistics, we analyze the privacy loss of the two algorithms and compare them. The key findings in this section are:

- SafeTab[Geometric] ensures pure differential privacy, and for the accuracy requirements specified by stakeholders in the US Census Bureau the total privacy loss is bounded by $\epsilon = 15.3$.
- SafeTab[Geometric] can also be analyzed under Renyi-DP to provide an approximate (ϵ, δ)-differential privacy guarantee. For the accuracy requirements specified by stakeholders in the US Census Bureau, the total privacy loss of SafeTab[Geometric] is bounded by $(\epsilon = 13.2, \delta = 10^{-10})$
- SafeTab[Discrete Gaussian] ensures zCDP which can be converted to an approximate (ϵ, δ)-differential privacy guarantee. For the accuracy requirements specified by stakeholders in the US Census Bureau, the total privacy loss of SafeTab[Discrete Gaussian] is bounded by $(\epsilon = 12.2, \delta = 10^{-10})$.

Finally, we considered the overall privacy loss that might result when the margin of error (MOE) is increased slightly for (Nation, detailed) and (State, detailed) population group levels from 6 to 7, 8, 9, 10 and 11. We observed that increasing the MOE results in a significant improvement on the overall privacy loss of both the Geometric (about 25% reduction) and Discrete Gaussian (about 30% reduction).

2 Problem & Desiderata

The SafeTab algorithm produces differentially private tables of statistics (counts) of demographic and housing characteristics of all persons in the US crossed with detailed races and tribes at varying levels of geography (national, state, county, AIANNH). In this section, we define relevant concepts, outline the statistics to be released, and then formulate the differentially private algorithm design problem.

2.1 Definitions

Every person (and household) resides in exactly one Census block that determines their geographic location. This Census block is contained in several *geographic entities* – e.g. LA county, the state of CA and the US. We assume there are hierarchical relationships between geographical entities.

Every person is also associated with one or more *race codes* and an *ethnicity code*. The maximum number of race codes a person can be associated with is called the *race multiplicity*. A *race (ethnicity) group* is a set of race (ethnicity) codes. An individual person is in a race group *Alone* if all race codes associated with that individual are contained in the race group. A record is in a race group *Alone or in Combination* if some race code associated with that record is contained in the race group.

A *characteristic iteration* is the combination of a race (or ethnicity) group, along with the specification of either Alone or Alone or in Combination. A Characteristic Iteration has a corresponding Characteristic Iteration Code. (E.g. Latin American Indian (6800-6999) Alone or in Combination is a characteristic iteration). Like geographical entities, characteristic iterations also have hierarchical relationships. One person may correspond to multiple characteristic iterations.

2.2 Detailed DHC (Race & Ethnicity) Data Product

The Detailed DHC (Race & Ethnicity) data product (DHC-RE) aims to tabulate statistics by population groups. A *population group* is a pair (g, c), where g is a geographic entity (e.g., the state of NC, or LA County) and c is a race or ethnicity characteristic iteration (e.g., Latin American Indian (6800-6999) Alone or in Combination).

Four different statistics are to be released for each population group. The first two statistics pertain to properties of persons:

(T1) Given as input a dataframe of all US persons and their attributes, output the total population associated with each population group.

(T2) Given as input a dataframe of all US persons and their attributes, output the Sex X Age 2-dimensional marginal for a subset of population groups (that are not marked as *TotalOnly*).

The other two statistics pertain to properties of households. Here, population groups are defined in terms of characteristic iterations and race/ethnicity properties of the householder.

(T3) Given as input a dataframe of all households in the US and their attributes, output a histogram of household types of households in each population group.

(T4) Given as input a dataframe of all households in the US and their attributes, output a histogram of tenure of households in each population group.

2.3 Private Release Problem

The release of statistical data products by the US Census Bureau about persons and households is regulated under Title 13 and any release of statistics about persons in the US must be afforded strong privacy protections [1]. Moreover, it has been demonstrated that legacy statistical disclosure limitation (SDL) techniques are vulnerable to attacks that can reconstruct the sensitive person records from aggregate statistics [3]. Hence, the US Census Bureau has decided that all statistics released as part of the 2020 Decennial Census (of which DHC-RE is a part) will be released using algorithms that satisfy modern privacy definitions like differential privacy [2].

In this paper, we describe SafeTab, a differentially private algorithm for releasing the statistics that make up the DHC-RE data product. **In this document, we focus on the algorithm for releasing tabulations T1 and T2 that take as input the dataframe of persons in the US**.

Based on workshops and in-depth discussions[1] with users of the DHC-RE data product and US Census Bureau data stewards, the following desiderata were identified for the differential privacy algorithm:

- *Privacy:* The algorithm must ensure end-to-end differential privacy with respect to (the addition/removal of) every person in the US.
- *Population Groups:* The algorithm must release statistics for a predefined set of race and ethnicity characteristic iterations and following geographies: national level (US), 50 states + DC, counties within the 50 states + DC and all areas designated as American Indian Alaska Native and Native Hawaian (AIANNH) areas.
- *Adaptivity:* The algorithm may adaptively choose the granularity at which Sex X Age statistics are released. For instance, for population groups with a few people the Sex X Age histogram may only have 4 buckets of age, while for population groups with many people a more detailed histogram may be released.
- *Accuracy:* Accuracy levels were specified for population groups in terms of the margin of error (MOE) in output counts. Different population groups had different MOEs specified (described later in the paper in Table 2).
- *Integrality:* The output statistics must be integral.
- *No Consistency:* The SafeTab differential privacy algorithm was not required to ensure consistency of any form. That is, different counts output by the systems need not be consistent with each other (e.g., the number of people of a certain characteristic iteration in the US need not equal the sum of the population counts for the same characteristic iterations across all states). The counts may also be negative. A separate statistical modeling algorithm is to be designed to ensure non-negativity of counts and certain forms of consistency – its description is out of scope for this paper.

In the rest of the paper, we describe the SafeTab differential privacy algorithms and analyze bounds on the privacy loss achievable while satisfying the constraints mentioned above.

[1]The methodology and tools used to elicit preferences of users on the statistics to be released, privacy parameters, accuracy constraints, etc. is out of scope for this paper and will be the focus of a separate paper.

3 SafeTab Algorithm

SafeTab-P is a privacy algorithm for releasing detailed race and ethnicity statistics from the 2020 Decennial Census. The algorithm must accommodate the release of tabulations for total counts by detailed race and ethnicity and tabulations for sex by age counts by detailed race and ethnicity. The algorithm acts on a private dataframe derived from the 2020 Decennial Census. There is a row for each person in the US with attributes for which census block the individual resides in, race and ethnicity codes, sex, and age.

In this section we present an analysis of a simplified version of the algorithm. In particular, SafeTab produces tabulations at the level of a *population group*. In reality, a population group is a geographic entity (e.g. a specific county) and a characteristic iteration code (see Section 2 for more details). Records are associated with population groups via algorithms that map their block id to geographic entities, and their race and ethnicity codes to iteration codes. Additionally, population groups are split into levels (both geography levels and iteration levels) with distinct privacy loss budgets. For the purposes of this section, we assume the following model for population groups:

- SafeTab should produce tabulations on sets of population groups $\mathcal{P}_1, \ldots, \mathcal{P}_\omega$, which we call *population group levels*. That is, it should produce a tabulation for each population group $P \in \mathcal{P}_i$ for $1 \leq i \leq \omega$.
- SafeTab is provided privacy loss budgets for each population group level $\rho_1, \ldots, \rho_\omega$ with ρ_i corresponding to the budget for population group level $\mathcal{P}_i$.
- For each $\mathcal{P}_i$, we assume we have a function $g_i : \mathcal{I} \to 2^{\mathcal{P}_i}$, where $\mathcal{I}$ is the domain of records in the private dataframe. That is, g_i maps a record to the subset of population groups at level i to which it belongs.
- We assume the stability of g_i, denoted by $\Delta(g_i)$ is known. The stability is defined as $\Delta(g_i) = \max_{r \in \mathcal{I}} |g_i(r)|$.

The main algorithm is presented in Algorithm 1. This algorithm proceeds by looping over the population group levels. For each population group level, we apply g_i to the dataframe to map each record to the set of population groups it is associated with. Then for each population group in the level, we call the tabulation function TABULATEPOPULATIONGROUP, passing in a dataframe containing just the records in that population group.

The pseudocode for the procedure TABULATEPOPULATIONGROUP is given in Algorithm 2. This code tabulates a single population group. Population groups are characterized based on the tabulation we would like to compute. In particular, we assume we are given a set $TotalOnly$ of population groups for which only the size of the population group should be tabulated. We check whether the given group is a member of this set. If it is, we call the NOISYCOUNT function on the population group, which tabulates a noisy count of the size of the group. Otherwise, we use a two stage algorithm. We first compute a noisy count of the group using NOISYCOUNT, but using only a fraction (denoted γ) of the available privacy loss budget. Next, we compare this noisy count against a set of given thresholds, denoted Θ_1, Θ_2, and Θ_3. Depending on which thresholds the noisy count exceeds, we compute sex by age noisy counts with a varying degree of age bin sizes. Age bins are coarser for smaller noisy counts. These sex by age counts are also computed by NOISYCOUNT using the remaining privacy loss budget.

The pseudocode for the procedure NOISYCOUNT is given in Algorithm 3. This procedure computes the number of rows in the dataframe and adds noise from either the discrete Gaussian

Notation	Description
ω	the number of population group levels
$\mathcal{P}_i$	population group level i
ρ_i	the privacy loss budget allocated to population group level i
g_i	a function mapping records to the set of population groups in $\mathcal{P}_i$ to which the record belongs
$\Delta(g_i)$	$\max_{r \in \mathcal{I}} \lvert g_i(r) \rvert$

Table 1: A summary of the notation used in Section 3

Algorithm 1 The main SafeTab-P [Γ] algorithm.

Input: df: private dataframe with attributes [BlockID, RaceEth, Sex, Age] and one row for each person in the US

Input: Γ: Noise Mechanism that is either Geometric Mechanism or Discrete Gaussian Mechanism

Input: $\{\rho_i\}_{i \in [1,\omega]}$: Privacy parameters for each population group level $i \in [1, \omega]$.

Input: γ: The fraction of the privacy loss budget to be used in Stage 1 of the two stage tabulation algorithm.

procedure SAFETAB-P(df, Γ, $\{\rho_i\}$, γ)
  for $i \in [1, \omega]$ **do**
    $df_i \leftarrow df$.flatmap(g_i); ▷ df_i has schema [PopGroup, Sex, Age]
    $s \leftarrow \Delta(g_i)$ ▷ 1 row in df may result in $\leq s$ rows in df_i
    for $P \in \mathcal{P}_i$ **do**
      $df_P \leftarrow df_i$.filter(PopGroup $== P$)
      TABULATEPOPULATIONGROUP(df_p, P, Γ, ρ_i/s, γ)
    end for
  end for
end procedure

distribution (see Algorithm 5) or the two-sided geometric distribution (see Algorithm 4), depending on the value of Γ. The parameter ρ can be interpreted as the pure DP loss (when Γ is *Geometric*) or as the zCDP loss (when Γ is *Discrete Gaussian*).

The notation used in this section and the algorithm pseudocode is summarized in Table 1.

Algorithm 2 Subroutine of SafeTab-P to tabulate statistics for a single population group.

Input: df: a private dataframe with attributes [BlockID, RaceEth, Sex, Age]. This dataframe should contain the records in the population group.
Input: P: The population group.
Input: Γ: Noise Mechanism that is either Geometric Mechanism or Discrete Gaussian Mechanism
Input: ρ: Privacy loss budget for this subroutine.
Input: γ: Fraction of ρ used in the adaptive algorithm

procedure TABULATEPOPULATIONGROUP($df, P, \Gamma, \rho, \gamma$)
  if $P \in$ TotalOnly **then**
    // For TotalOnly population groups, only report noisy total counts
    Output NOISYCOUNT(df, Γ, ρ)

  else
    // For the rest of the population groups, adaptively choose the statistics released
    // based on the noisy total count of the population group.
    // Step 1: Compute the noisy total count using $\rho\gamma$ privacy loss budget
    total $\leftarrow$ NOISYCOUNT($df, \Gamma, \gamma\rho$) ▷ Compute noisy total

    // Step 2: Release statistics based on the noisy count with $\rho(1-\gamma)$ privacy loss budget
    if total $< \theta_1$ **then**
      Output NOISYCOUNT(df, Γ, $(1-\gamma)\rho$) ▷ Output the total
    else if total $< \theta_2$ **then**
      for df_group $\in$ df.map(Age $\rightarrow$ Age4).groupby(Sex, Age4) **do**
        Output NOISYCOUNT(df_group, Γ, $(1-\gamma)\rho$) ▷ Sex X Age4 marginal
      end for
    else if total $< \theta_3$ **then**
      for df_group $\in$ df.map(Age $\rightarrow$ Age9).groupby(Sex, Age9) **do**
        Output NOISYCOUNT(df_group, Γ, $(1-\gamma)\rho$) ▷ Sex X Age9 marginal
      end for
    else
      for df_group $\in$ df.map(Age $\rightarrow$ Age23).groupby(Sex, Age23) **do**
        Output NOISYCOUNT(df_group, Γ, $(1-\gamma)\rho$) ▷ Sex X Age23 marginal
      end for
    end if
  end if
end procedure

Algorithm 3 Noisy Count Mechanism

Input: df: The dataframe.
Input: Γ: The noise mechanism. This should be either Geometric or Discrete Gaussian.
Input: ρ: The privacy loss budget for this subroutine. Its interpretation depends on the chosen noise mechanism.

procedure NOISYCOUNT(df, Γ, ρ)
  if Γ is Geometric **then**
    // ρ is the pure-DP privacy loss parameter
    return BASEGEOMETRIC(df.count(), ρ)
  else if Γ is Discrete Gaussian **then**
    // ρ is the zCDP privacy loss parameter
    return BASEDISCRETEGAUSSIAN(df.count(), ρ)
  end if
end procedure

Algorithm 4 The base geometric mechanism.

Input: c: An integer.
Input: ϵ: The desired privacy loss parameter.

procedure BASEGEOMETRIC(c, ϵ)
  $y \leftarrow \mathcal{L}_{\mathbb{Z}}\left(\frac{1}{\epsilon}\right)$
  return $c + y$
end procedure

Algorithm 5 The base discrete Gaussian mechanism.

Input: c: An integer.
Input: ρ: The desired privacy loss parameter.

procedure BASEDISCRETEGAUSSIAN(c, ρ)
  $y \leftarrow \mathcal{N}_{\mathbb{Z}}\left(\frac{1}{2\rho}\right)$
  return $c + y$
end procedure

4 Privacy Preliminaries

In this section, we give necessary background on differential privacy and related definitions of privacy. In particular, we will analyze SafeTab using pure and approximate differential privacy, and zCDP, as well as with two different basic noise mechanisms, the geometric mechanism and the discrete Gaussian mechanism.

4.1 Privacy definitions

Definition 1 (Neighboring Databases)**.** Let x, x' be databases represented as multisets of tuples. We say that x and x' are *neighbors* if their symmetric difference is 1.

We first define diffential privacy, the most common formal privacy definition.

Definition 2. An algorithm $M : \mathcal{X} \to \mathcal{Y}$ satisfies (ϵ, δ)-differential privacy if for all neighboring databases x, x' and all output $y \in \mathcal{Y}$,

$$P[M(x) = y] \leq e^{\epsilon} P[M(x') = y] + \delta \tag{1}$$

When a mechanism satisfies differential privacy with $\delta = 0$, we say that the mechanism satisfies *pure differential privacy*, and when $\delta > 0$ we say the mechanism satisfies *approximate differential privacy*.

We next define zCDP, which bounds the *Rényi divergence* between between the distributions of a mechanism run on neighboring databases.

Definition 3. The *Rényi divergence of order* α between distribution P and distribution Q, denoted $D_\alpha(P\|Q)$ is defined as

$$D_\alpha(P\|Q) = \frac{1}{\alpha - 1} \log \left(\mathop{\mathbb{E}}_{x \sim P} \left[\left(\frac{P(x)}{Q(x)} \right)^{\alpha - 1} \right] \right) \tag{2}$$

Definition 4. (zCDP [4]) An algorithm $M : \mathcal{X} \to \mathcal{Y}$ satisfies ρ-zCDP if for all neighboring $x, x' \in \mathcal{X}$ and for all $\alpha \in (1, \infty)$,

$$D_\alpha(M(x)\|M(x')) \leq \rho\alpha. \tag{3}$$

Finally, we define Rényi differential privacy (RDP). RDP is similar to zCDP except that it (1) bounds the Rényi divergence of each order separately, and (2) allows for an arbitrary bound on the divergence, rather than requiring a bound that is linear in α.

Definition 5 (RDP [7])**.** An algorithm $M : \mathcal{X} \to \mathcal{Y}$ satisfies (α, ϵ)-Rényi differential privacy ((α, ϵ)-RDP) if for all neighboring $x, x' \in \mathcal{X}$,

$$D_\alpha(M(x)\|M(x')) \leq \epsilon. \tag{4}$$

4.2 Composition

One of the most useful and important properties of privacy definitions is their behaviour under composition. In this section, we state composition results for pure differential privacy, approximate differential privacy, zCDP, and RDP. There are two types of composition we are interested in – sequential composition and parallel composition. We first state the sequential composition results.

Lemma 1. *(Adaptive sequential composition of pure differential privacy) Let $M_1 : \mathcal{X} \to \mathcal{Y}$ and $M_2 : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be mechanisms satisfying ϵ_1-differential privacy and ϵ_2-differential privacy respectively. Let $M_3(x) = M_2(x, M_1(x))$. Then M_3 satisfies $(\epsilon_1 + \epsilon_2)$-differential privacy.*

Lemma 2. *(Adaptive sequential composition of zCDP [4]) Let $M_1 : \mathcal{X} \to \mathcal{Y}$ and $M_2 : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be mechanisms satisfying ρ_1-zCDP and ρ_2-zCDP respectively. Let $M_3(x) = M_2(x, M_1(x))$. Then M_3 satisfies $(\rho_1 + \rho_2)$-zCDP.*

Lemma 3. *(Adaptive sequential composition of RDP [7]) Let $M_1 : \mathcal{X} \to \mathcal{Y}$ and $M_2 : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be mechanisms satisfying (α, ϵ_1)-RDP and (α, ϵ_2)-RDP respectively. Let $M_3(x) = M_2(x, M_1(x))$. Then M_3 satisfies $(\alpha, \epsilon_1 + \epsilon_2)$-RDP.*

In Section 5.2 we state and prove generalized parallel composition lemmas for our privacy definitions.

4.3 Converting zCDP and RDP to approximate differential privacy

Lemma 4. *([5]) Let $M : \mathcal{X} \to \mathcal{Y}$ be a randomized algorithm and let $\alpha \in (1, \infty)$ and $\epsilon > 0$. Suppose $D_\alpha(M(x) \| M(x')) \leq \tau$ for all neighboring x, x'. Then M satisfies (ϵ, δ)-differential privacy where*

$$\delta = \frac{\exp((\alpha - 1)(\tau - \epsilon))}{\alpha - 1} \left(1 - \frac{1}{\alpha}\right)^\alpha, \tag{5}$$

or equivalently,

$$\epsilon = \tau + \frac{\log(1/\delta) + (\alpha - 1)\log(1 - 1/\alpha) - \log(\alpha)}{\alpha - 1}. \tag{6}$$

Lemma 4 can be used to convert both RDP and zCDP guarantees to approximate differential privacy guarantees. For RDP, the conversion follows immediately from the lemma (τ in the lemma above the is the RDP ϵ parameter, which is different from the approximate differential privacy ϵ parameter). Converting zCDP to approximate differential privacy using Lemma 4 requires minimizing over all possible values of α to find the tightest guarantee. That is, if M satisfies ρ-zCDP then it satisfies (ϵ, δ)-differential privacy where

$$\epsilon = \inf_{\alpha \in (1, \infty)} \left[\rho\alpha + \frac{\log(1/\delta) + (\alpha - 1)\log(1 - 1/\alpha) - \log(\alpha)}{\alpha - 1} \right]. \tag{7}$$

Computing this expression for ϵ can be done experimentally, but cannot easily be done analytically. Because of this, we also give a looser conversion from zCDP to approximate differential privacy [4].

Lemma 5. *([4]) Let $M : \mathcal{X} \to \mathcal{Y}$ be a randomized algorithm satisfying ρ-zCDP. Then for all $\delta > 0$, M satisfies (ϵ, δ)-differential privacy where*

$$\epsilon = \rho + \sqrt{4\rho \log(1/\delta)}. \tag{8}$$

4.4 Base Mechanisms

Definition 6. The discrete gaussian distribution $\mathcal{N}_\mathbb{Z}(\sigma^2)$ centered at 0 is

$$\forall x \in \mathbb{Z}, \quad \Pr[X = x] = \frac{e^{-x^2/2\sigma^2}}{\sum_{y \in \mathbb{Z}} e^{-y^2/2\sigma^2}}. \tag{9}$$

Definition 7. The two sided geometric (or discrete Laplace) distribution $\mathcal{L}_{\mathbb{Z}}(b)$ centered at 0 is

$$\forall x \in \mathbb{Z}, \quad \Pr[X = x] = \frac{e^{1/b} - 1}{e^{1/b} + 1} \cdot e^{-|x|/b}. \tag{10}$$

Lemma 6. *Let $q : \mathcal{X} \to \mathbb{R}$. Then* BASEGEOMETRIC$(q(x), \epsilon)$ *from Algorithm 4 satisfies ϵ-differential privacy with respect to x.*

Lemma 7. *[5] Let $q : \mathcal{X} \to \mathbb{R}$. Then* BASEDISCRETEGAUSSIAN$(q(x), \rho)$ *from Algorithm 5 satisfies ρ-zCDP with respect to x.*

Note that

- BASEDISCRETEGAUSSIAN does not satisfy pure differential privacy for any value of ϵ.
- Any ϵ-differentially private algorithm also satisfies $(\frac{1}{2}\epsilon^2)$-zCDP. There, BASEGEOMETRIC$(q(x), \epsilon)$ satisfies $(\frac{1}{2}\epsilon^2)$-zCDP.
- Analyzing BASEGEOMETRIC-based algorithms using $zCDP$ leads to looser bounds than using an RDP analysis. Therefore, in this paper we only consider an RDP analysis. The RDP guarantee for BASEGEOMETRIC is given in Corollary 1.

5 Novel Privacy Results

In this section, we give some useful privacy results necessary for deriving the privacy losses of SafeTab. To the best of our knowledge, these results are novel.

5.1 Rényi parameters for base geometric

We show that the base geometric mechanism satisfies Rényi differential privacy and give the Rènyi privacy parameters for which this is true. This will be useful for finding an approximate differential privacy guarantee for the composition of the geometric mechanism. Lemma 8 is an analogue of Proposition 6 in [7], and the proof is similar.

Lemma 8. *For any $\alpha > 1$ and $b > 0$,*

$$D_\alpha(\mathcal{L}_{\mathbb{Z}}(b) \| (\mathcal{L}_{\mathbb{Z}}(b) + 1)) = \frac{1}{\alpha - 1} \log \left[\frac{e^{1/b} - 1}{e^{1/b} + 1} \left(\frac{2\alpha b}{2\alpha - 1} e^{(\alpha-1)/b} + \frac{2\alpha b - 2b}{2\alpha - 1} e^{-\alpha/b} \right) \right] \tag{11}$$

Proof. Let P and Q be distributions with densities $p(x)$ and $q(x)$ respectively. Then,

$$D_\alpha(P \| Q) = \frac{1}{\alpha - 1} \log \int_{-\infty}^{\infty} p(x)^\alpha q(x)^{1-\alpha} \, \mathrm{d}x. \tag{12}$$

For us,

$$p(x) = \frac{e^{1/b} - 1}{e^{1/b} + 1} \cdot e^{-|x|/b}, \quad q(x) = \frac{e^{1/b} - 1}{e^{1/b} + 1} \cdot e^{-|x-1|/b}. \tag{13}$$

We evaluate the integral separately on $(-\infty, 0)$, $(0, 1)$, and $(1, \infty)$. That is,

$$\int_{-\infty}^{\infty} p(x)^{\alpha} q(x)^{1-\alpha} \, \mathrm{d}x = \frac{e^{1/b} - 1}{e^{1/b} + 1} \cdot \left[\int_{-\infty}^{0} \exp\left(\frac{\alpha x}{b} + \frac{(1-\alpha)(x-1)}{b} \right) \mathrm{d}x \right. \tag{14}$$

$$\left. + \int_{0}^{1} \exp\left(-\frac{\alpha x}{b} + \frac{(1-\alpha)(x-1)}{b} \right) \mathrm{d}x \right.$$

$$\left. + \int_{1}^{\infty} \exp\left(-\frac{\alpha x}{b} - \frac{(1-\alpha)(x-1)}{b} \right) \mathrm{d}x \right]$$

$$= \frac{e^{1/b} - 1}{e^{1/b} + 1} \cdot \left[b e^{(\alpha-1)/b} + \frac{b}{2\alpha - 1} \left(e^{(\alpha-1)/b} - e^{-\alpha/b} \right) + b e^{-\alpha/b} \right] \tag{15}$$

$$= \frac{e^{1/b} - 1}{e^{1/b} + 1} \left(\frac{2\alpha b}{2\alpha - 1} e^{(\alpha-1)/b} + \frac{2\alpha b - 2b}{2\alpha - 1} e^{-\alpha/b} \right). \tag{16}$$

□

In the corollary below, we use the notation τ to denote the privacy loss instead of the usual notation ϵ in order to differentiate it from the pure DP loss that denote by ϵ.

Corollary 1. *Let $q : \mathcal{X} \to \mathbb{R}$. Then* BASEGEOMETRIC$(q(x), \epsilon)$ *satisfies (α, τ)-RDP privacy with respect to x, where*

$$\tau = \frac{1}{\alpha - 1} \log \left[\frac{e^{\epsilon} - 1}{e^{\epsilon} + 1} \left(\frac{2\alpha}{\epsilon(2\alpha - 1)} e^{(\alpha-1)\epsilon} + \frac{2(\alpha - 1)}{\epsilon(2\alpha - 1)} e^{-\alpha\epsilon} \right) \right]. \tag{17}$$

5.2 Generalized parallel composition lemmas

In this section, we give generalized parallel composition lemmas for pure DP, zCDP, and RDP. The statements we give are generalizations of the standard statements of parallel composition.

Let the *maximum degree* of a set family $F = \{S_i\}$, $S_i \subseteq S$ be the maximum number of sets containing any fixed element of S. That is,

$$degree(F) = max_{s \in S} |\{S_i \in F | s \in S_i\}| \tag{18}$$

Lemma 9. *Let $F = \{S_1, ..., S_k\}$ be family of subsets of the input domain with maximum degree z. Let $M_1, \ldots, M_k$ each provide ϵ-differential privacy. Then the mechanism $M(x) = (M_1(x \cap S_1), \ldots, M_k(x \cap S_k))$ provides $(z \cdot \epsilon)$-differential privacy.*

Proof. Suppose x and x' are neighbors, and let r be the (only) record in their symmetric difference. Let $i_1, \ldots, i_j$ be the indices of the sets in F containing r. $j \leq z$ since the maximum degree of F is

z. Then for any y,

$$\Pr[M(x) = y] = \prod_{i=1}^{k} \Pr[M_i(x \cap S_i) = y_i] \tag{19}$$

$$= \prod_{i \in \{i_1,\ldots,i_j\}} \Pr[M_i(x \cap S_i) = y_i] \prod_{i \notin \{i_1,\ldots,i_j\}} \Pr[M_i(x' \cap S_i) = y_i] \tag{20}$$

$$\leq \prod_{i \in \{i_1,\ldots,i_j\}} \left[e^{\epsilon} \Pr[M_i(x' \cap S_i) = y_i]\right] \prod_{i \notin \{i_1,\ldots,i_j\}} \Pr[M_i(x' \cap S_i) = y_i] \tag{21}$$

$$= e^{j\epsilon} \cdot \prod_{i=1}^{k} \Pr[M_i(x' \cap S_i) = y_i] \tag{22}$$

$$= e^{j\epsilon} \cdot \Pr[M(x') = y]. \tag{23}$$

□

We can give an analogous lemma for zCDP.

Lemma 10. *Let $F = \{S_1, ..., S_k\}$ be family of subsets of the input domain with maximum degree z. Let $M_1, \ldots, M_k$ each provide ρ-zCDP. Then the mechanism $M(x) = (M_1(x \cap S_1), \ldots, M_k(x \cap S_k))$ provides $(z \cdot \rho)$-zCDP.*

The proof of 10 requires the following property on the Rényi divergence, given in Lemma 2.2 of [4].

Lemma 11. *([4]) Let P_Ω and Q_Ω be distributions on Ω, and P_Θ and Q_Θ be distributions on Θ. Let $P = P_\Omega P_\Theta$ and $Q = Q_\Omega Q_\Theta$. Then*

$$D_\alpha(P \| Q) = D_\alpha(P_\Omega || Q_\Omega) + D_\alpha(P_\Theta || Q_\Theta) \tag{24}$$

With this, we can prove Lemma 10.

Proof of Lemma 10. Suppose x and x' are neighbors, and let r be the (only) record in their symmetric difference. Let $i_1, \ldots, i_j$ be the indices of the sets in F containing r. $j \leq z$ since the maximum degree of F is z.

$$D_\alpha(M(x) \| M(x')) = \sum_{i=1}^{k} D_\alpha(M_i(x \cap S_i) \| M_i(x' \cap S_i)) \tag{25}$$

$$= \sum_{i \in \{i_1,\ldots,i_j\}} D_\alpha(M_i(x \cap S_i) \| M_i(x' \cap S_i)) \tag{26}$$

$$\leq \sum_{i \in \{i_1,\ldots,i_j\}} \alpha \cdot \rho \tag{27}$$

$$\leq \alpha \cdot (z \cdot \rho). \tag{28}$$

□

Finally, we give a generalized parallel composition lemma for RDP. The proof is nearly identical to the proof of Lemma 10 so we omit it.

Lemma 12. *Let $F = \{S_1, ..., S_k\}$ be family of subsets of the input domain with maximum degree z. Let $M_1, \ldots, M_k$ each provide (α, ϵ)-RDP. Then the mechanism $M(x) = (M_1(x \cap S_1), \ldots, M_k(x \cap S_k))$ provides $(\alpha, z \cdot \epsilon)$-RDP.*

6 SafeTab[Geometric] Privacy and Error Analysis

In this section, we give two privacy analyses for the SafeTab[Geometric] algorithm, as well as an analysis of the error of the algorithm. Both the privacy analyses in this section, as well as the privacy analysis in Section 7, follow a very similar recipe. Each result follows from the privacy properties of the base mechanism, combined with the composition rules given in sections 4 and 5. Because the composition results for the different privacy definitions are essentially the same, the privacy analysis proofs are all very similar. For completeness, we give all the privacy proofs.

Note that it is also possible to give a zCDP analysis of SafeTab[Geometric]. However, this analysis is not as tight as RDP analysis so we omit it.

6.1 Pure-DP Privacy Analysis

In this section, show that the SafeTab[Geometric] algorithm presented in Section 3 satisfies pure differential privacy.

Note that we chose to use ρ as a parameter in the SafeTab algorithm description. When analyzing the geometric version of the algorithm under pure DP, the parameter ρ corresponds to the pure dp loss, which is generally denoted as ϵ.

Theorem 1. *Let $\rho_{total} = \sum_{i=1}^{\omega} \rho_i$. Algorithm 1 satisfies ρ_{total}-differential privacy when Γ is the Geometric Mechanism.*

Proof. The proof follows from a combination of composition rules along with the fact that the base mechanism, BASEGEOMETRIC satisfies pure dp.

First, we claim that the procedure NOISYCOUNT where Γ = *Geometric* satisfies ρ-differential privacy, where ρ is the privacy parameter input to NOISYCOUNT. This follows directly from Lemma 6.

Next, we claim that the procedure TABULATEPOPULATIONGROUP in Algorithm 2 satisfies ρ-differential privacy with respect to the input dataframe, where ρ is the privacy parameter input to the procedure. Note that TABULATEPOPULATIONGROUP actually uses one of two algorithms depending on whether the population group is in the set TotalOnly. We consider each of these algorithms.

Case 1: $P \in$ TotalOnly. In this case the procedure simply calls NOISYCOUNT, which satisfies ρ-differential privacy.

Case 2: $P \notin$ TotalOnly. In this case, the procedure can be decomposed into two parts. First, we call NOISYCOUNT with a budget of $\gamma\rho$. Then, we use the result to group the data by sex and age, and for each group we make a call to NOISYCOUNT with a budget of $(1-\gamma)\rho$. The composition of the calls on all the groups satisfies $(1-\gamma)\rho$ by Lemma 9. The (adaptive) composition of the two parts has total privacy loss ρ by Lemma 1.

Next, we claim that the ith loop of the **for** loop on line 2 of Algorithm 1 satisfies ρ_i-differential privacy. By the definition of s, any particular record can appear in the input (df_P) of at most s calls to TABULATEPOPULATIONGROUP. Therefore, by Lemma 9, the total privacy loss of the loop is s times the privacy loss of TABULATEPOPULATIONGROUP, i.e. $s \cdot \frac{\rho_i}{s} = \rho_i$.

Finally, the overall algorithm satisfies $(\sum_{i=1}^{\omega} \rho_i)$-differential privacy by Lemma 1. □

6.1.1 A note on the generalized parallel composition lemma

Rather than using the generalized parallel composition lemma (Lemma 9) to analyze the algorithm, we could have used the popular *stability*-based accounting method [6]. This is an approach

for calculating the sensitivity of a query made up of multiple database transformations followed by a measurement. This method gives a tight analysis under differential privacy using the geometric mechanism because combining a sensitivity x query and a sensitivity y query into a single query (outputting a vector) has sensitivity $x + y$. Therefore, analyzing the algorithm as a few high sensitivity queries (stability analysis) is equivalent to analyzing it as the the composition of many low sensitivity queries.

On the other hand, adding noise from the discrete Gaussian distribution to satisfy zCDP requires the scale of the noise to be proportional to the square of the sensitivity, rather than the sensitivity. This means that a single query with sensitivity $x + y$ will require significantly more noise than the composition of mechanisms that answer sensitivity x and sensitivity y queries respectively. Therefore, we can get a tighter analysis by analyzing the algorithm as the composition of many low sensitivity queries (sensitivity 1 in the case of SafeTab) rather than fewer high sensitivity queries. This requires using the generalized parallel composition instead of stability accounting.

6.2 RDP Privacy Analysis

In this section, we show that the SafeTab[Geometric] algorithm presented in Section 3 satisfies RDP.

Theorem 2. *Let $\tau(\alpha, \epsilon)$ denote the RDP privacy bound for* BASEGEOMETRIC *given in Corollary 1. That is,*

$$\tau(\alpha, \epsilon) = \frac{1}{\alpha - 1} \log \left[\frac{e^\epsilon - 1}{e^\epsilon + 1} \left(\frac{2\alpha}{\epsilon(2\alpha - 1)} e^{(\alpha-1)\epsilon} + \frac{2(\alpha - 1)}{\epsilon(2\alpha - 1)} e^{-\alpha\epsilon} \right) \right]. \tag{29}$$

Algorithm 1 satisfies

$$\left(\alpha, \sum_{i=1}^{\omega} \left[s \cdot \max \left[\tau \left(\alpha, \frac{\gamma\rho_i}{s} \right) + \tau \left(\alpha, \frac{(1-\gamma)\rho_i}{s} \right), \tau \left(\alpha, \frac{\rho_i}{s} \right) \right] \right] \right) \text{-RDP} \tag{30}$$

when Γ is the Geometric Mechanism.

Proof. The proof follows from a combination of composition rules along with the fact that the base mechanism, BASEGEOMETRIC satisfies $(\alpha, \tau(\alpha, \epsilon))$-RDP.

First, we claim that the procedure NOISYCOUNT where $\Gamma = \textit{Geometric}$ satisfies $(\alpha, \tau(\alpha, \rho))$-RDP, where ρ is the privacy parameter input to NOISYCOUNT. This follows directly from Corollary 1.

Next, we claim that the procedure TABULATEPOPULATIONGROUP in Algorithm 2 satisfies $(\alpha, \max(\tau(\alpha, \gamma\rho) + \tau(\alpha, (1-\gamma)\rho), \tau(\alpha, \rho)))$-RDP with respect to the input dataframe, where ρ is the privacy parameter input to the procedure. Note that TABULATEPOPULATIONGROUP actually uses one of two algorithms depending on whether the population group is in the set TotalOnly. We consider each of these algorithms.

Case 1: $P \in$ TotalOnly. In this case the procedure simply calls NOISYCOUNT, which satisfies $(\alpha, \tau(\alpha, \rho))$-RDP.

Case 2: $P \notin$ TotalOnly. In this case, the procedure can be decomposed into two parts. First, we call NOISYCOUNT with a budget of $\gamma\rho$. Then, we use the result to group the data by sex and age, and for each group we make a call to NOISYCOUNT with a budget of $(1-\gamma)\rho$. The composition of the calls on all the groups satisfies $(\alpha, \tau(\alpha, (1-\gamma)\rho))$-RDP by Lemma 12. The (adaptive) composition of the two parts satisfies $(\alpha, \tau(\alpha, \gamma\rho) + \tau(\alpha, (1-\gamma)\rho))$-RDP by Lemma 3.

Next, we claim that the ith loop of the **for** loop on line 2 of Algorithm 1 satisfies

$$\left(\alpha, s \cdot \max \left[\tau \left(\alpha, \frac{\gamma\rho_i}{s} \right) + \tau \left(\alpha, \frac{(1-\gamma)\rho_i}{s} \right), \tau \left(\alpha, \frac{\rho_i}{s} \right) \right] \right) \text{-RDP}. \tag{31}$$

By the definition of s, any particular record can appear in the input (df_P) of at most s calls to TABULATEPOPULATIONGROUP. Therefore, by Lemma 12, the total privacy loss of the loop is s times the privacy loss of TABULATEPOPULATIONGROUP.

Finally, the overall algorithm satisfies

$$\left(\alpha, \sum_{i=1}^{\omega}\left[s \cdot \max\left[\tau\left(\alpha, \frac{\gamma\rho_i}{s}\right) + \tau\left(\alpha, \frac{(1-\gamma)\rho_i}{s}\right), \tau\left(\alpha, \frac{\rho_i}{s}\right)\right]\right]\right) \text{-RDP} \tag{32}$$

by Lemma 3.

□

6.3 Error Bounds

We next examine the utility of Algorithm 1 with base geometric mechanism. We first restate a portion of Lemma 30 from [5]:

Lemma 13. *Let $b > 0$ and let $Y \leftarrow \mathcal{L}_{\mathbb{Z}}(b)$. For all $y \in \mathbb{R}$,*

$$\Pr[Y \geq y] = \Pr[Y \leq -y] \leq \frac{e^{-\frac{\lceil y \rceil}{b}}}{1 + e^{-\frac{1}{b}}}. \tag{33}$$

Hence for all $y \in \mathbb{R}$,

$$P[Y > y] = P[Y < -y] \leq \frac{e^{-\frac{\lfloor y \rfloor}{b}}}{1 + e^{-\frac{1}{b}}} - \frac{e^{1/b} - 1}{e^{1/b} + 1} \cdot e^{-\lfloor y \rfloor / b} = \frac{e^{-\lfloor y \rfloor / b}}{1 + e^{1/b}}. \tag{34}$$

It follows that,

$$Y \in \left[-\left\lfloor b \ln\left(\frac{2}{p\left(1 + e^{\frac{1}{b}}\right)}\right)\right\rfloor, \left\lfloor b \ln\left(\frac{2}{p\left(1 + e^{\frac{1}{b}}\right)}\right)\right\rfloor\right]. \tag{35}$$

with probability $1 - p$. Hence, the margin of error of a 95% confidence interval is

$$\left\lfloor b \ln\left(\frac{40}{1 + e^{\frac{1}{b}}}\right)\right\rfloor. \tag{36}$$

Note that for a fixed integral 95% MOE, we have $b \in \left(\frac{MOE}{\ln 20}, \frac{MOE+1}{\ln 20}\right)$.

Corollary 2. *The base geometric mechanism run with parameter $\epsilon = \frac{\ln(20)}{\lfloor MOE \rfloor + 1}$ has a 95% margin on error of at most MOE.*

Applying the error bounds, the total count estimate of a TotalOnly population group in population group level i would have a margin of error of $\left\lfloor \frac{s}{\rho_i} \ln\left(\frac{40}{1+e^{\rho_i/s}}\right)\right\rfloor$. For a non-TotalOnly population group in level i, the margin of error in a single sex by age group is $\left\lfloor \frac{s}{(1-\gamma)\rho_i} \ln\left(\frac{40}{1+e^{(1-\gamma)\rho_i/s}}\right)\right\rfloor$.

7 SafeTab[Discrete Gaussian] Privacy Analysis

In this section, we give a zCDP analysis of the SafeTab[Discrete Gaussian] algorithm, as well as an analysis of the error of the algorithm. This analysis follows the same formula as the analyses in Section 6.

7.1 zCDP privacy analysis

In this section, show that the SafeTab[Discrete Gaussian] algorithm presented in Section 3 satisfies zero-concentrated differential privacy (zCDP).

Theorem 3. *Let $\rho_{total} = \sum_{i=1}^{\omega} \rho_i$. Algorithm 1 satisfies ρ_{total}-zCDP when Γ is the Discrete Gaussian Mechanism.*

Proof. The proof of Theorem 3 via the combination of composition rules along with the fact that the base mechanism, BASEDISCRETEGAUSSIAN.

First, we claim that the procedure NOISYCOUNT where $\Gamma =$ *Discrete Gaussian* satisfies ρ-zCDP, where ρ is the privacy parameter input to NOISYCOUNT. This follows directly from Lemma 7.

Next, we claim that the procedure TABULATEPOPULATIONGROUP in Algorithm 2 satisfies ρ-zCDP with respect to the input dataframe, where ρ is the privacy parameter input to the procedure. Note that TABULATEPOPULATIONGROUP actually uses one of two algorithms depending on whether the population group is in the set TotalOnly. We consider each of these algorithms.

Case 1: $P \in$ TotalOnly. In this case the procedure simply calls NOISYCOUNT, which satisfies ρ-zCDP.

Case 2: $P \notin$ TotalOnly. In this case, the procedure can be decomposed into two parts. First, we call NOISYCOUNT with a budget of $\gamma\rho$. Then, we use the result to group the data by sex and age, and for each group we make a call to NOISYCOUNT with a budget of $(1-\gamma)\rho$. The composition of the calls on all the groups satisfies $(1-\gamma)\rho$ by Lemma 10. The (adaptive) composition of the two parts has total privacy loss ρ by Lemma 2.

Next, we claim that the ith loop of the **for** loop on line 2 of Algorithm 1 satisfies ρ_i-zCDP. By the definition of s, any particular record can appear in the input (df_P) of at most s calls to TABULATEPOPULATIONGROUP. Therefore, by Lemma 10, the total privacy loss of the loop is s times the privacy loss of TABULATEPOPULATIONGROUP, i.e. $s \cdot \frac{\rho_i}{s} = \rho_i$.

Finally, the overall algorithm satisfies $(\sum_{i=1}^{\omega} \rho_i)$-zCDP by Lemma 2. □

7.2 Error bounds

We next examine the utility of Algorithm 1 with discrete Gaussian noise. We begin by stating a portion of Proposition 25 from [5].

Proposition 1 (Proposition 25 in [5])**.** *For all $m \in \mathbb{Z}$ with $m \geq 1$, and for all $\sigma \in \mathbb{R}$ with $\sigma > 0$, $\Pr[X \geq m]_{X \leftarrow \mathcal{N}_{\mathbb{Z}}(\sigma^2)} \leq \Pr[X \geq m-1]_{X \leftarrow \mathcal{N}(\sigma^2)}$.*

The following corollary is immediate.

Corollary 3. *For all $m, \sigma \in \mathbb{R}$ with $x \geq 1$ and $\sigma > 0$, $\Pr[X > x]_{X \leftarrow \mathcal{N}_{\mathbb{Z}}(\sigma^2)} \leq \Pr[X > \lfloor x \rfloor]_{X \leftarrow \mathcal{N}(\sigma^2)}$.*

Figure 2 of [5] provides an intuitive visualization of these tail bounds. It follows that $X \in [-\lfloor 1.96\sigma \rfloor, \lfloor 1.96\sigma \rfloor]$ with probability at least 95%. That is, the 95% margin of error is given by $\lfloor 1.96\sigma \rfloor$.

Hence for a population group in level i in the TotalOnly set, the margin of error in the directly computed total estimate from line 4 in Algorithm 2 is $\left\lfloor 1.96\sqrt{\frac{s}{2\rho_i}} \right\rfloor$. For the population groups in level i not in the TotalOnly set, the margin of error in a single sex by age group in Algorithm 2 is $\left\lfloor 1.96\sqrt{\frac{s}{2(1-\gamma)\rho_i}} \right\rfloor$.

Corollary 4. *The base discrete Gaussian mechanism run with $\rho = \frac{1.92}{\lfloor MOE \rfloor^2}$ has a 95% margin on error of at most MOE.*

Population Group Level	MOE Target	Geometric (ϵ)		Discrete Gaussian (ρ)	
		Step 2	Total	Step 2	Total
(Nation, Detailed)	6	3.84	4.27	0.481	0.534
(State, Detailed)	6	3.84	4.27	0.481	0.534
(County, Detailed)	11	2.24	2.49	0.143	0.159
(AIANNH, Detailed)	11	2.24	2.49	0.143	0.159
(Nation, Regional)	50	0.531	0.59	0.007	0.008
(State, Regional)	50	0.531	0.59	0.007	0.008
(County, Regional)	50	0.531	0.59	0.007	0.008

Table 2: MOE targets for the statistics released (in Step 2 of the adaptive algorithm) at different population group levels along with the corresponding privacy loss (ϵ-DP for Geometric and ρ-zCDP for Discrete Gaussian). The privacy loss is reported for the Step 2 (to match the MOE) as well as the total loss for that level. Step 2 loss is 90% of Total loss at each population group level. Note that the privacy losses reported here have already been aggregated over all the population groups at the given level, so the *Total* column represents the privacy loss input parameters of the SafeTab algorithm.

8 Comparing SafeTab[Geomteric] vs SafeTab[Discrete Gaussian]

In this section we compare SafeTab[Discrete Gaussian] and SafeTab[Geometric]. We set parameters to their desired production settings, and compare the approximate dp privacy loss of the algorithms using various analyses. In particular, we fix a target margin of error for both versions of the algorithm and compute the approximate differential privacy loss for a fixed δ.

8.1 Fixing algorithm parameters

To evaluate the privacy losses of the algorithm under approximate differential privacy, we set parameter values specified by the Census. Parameters are set as follows.

- The approximate differential privacy parameter δ is 10^{-10}.
- The number of population group levels is 7. These population groups are enumerated in the *Population Group Level* column of Table 2.
- The population group mapping function g_i has stability $\Delta(g_i) = 9$ for all i.
- The parameter γ is 0.1.
- The MOE targets are given in Table 2. Note that MOE targets are set for Step 2 of the 2-step algorithm (i.e. lines 14 to 25 of Algorithm 2).

We next convert the MOE targets into privacy parameter algorithm inputs as follows. First, we convert the MOE target for each base mechanism into the corresponding privacy parameter for the mechanism using the tail bounds presented in the previous section. That is, the parameter ρ for the base discrete Gaussian mechanism is given by

$$\rho = \frac{1.92}{\lfloor MOE \rfloor^2}, \tag{37}$$

where MOE is the target MOE for the base mechanism. See Section 7.2 for the details of the MOE derivation.

The parameter ϵ for the base geometric mechanism is

$$\epsilon = \frac{\ln(20)}{\lfloor MOE \rfloor + 1}. \tag{38}$$

See Section 6.3 for the details of the MOE derivation.

We next compute the parameter ρ that should be given as input to TABULATEPOPULATIONGROUP. Since our MOE target is for step 2 of the two step algorithm, this parameter ρ is given by $\rho = (1/\gamma)\rho_{base}$ where ρ_{base} is the privacy parameter we computed for the base mechanism using the MOE target.

Finally, we compute each of $\rho_1, \ldots, \rho_7$ by multiplying the ρ parameter from each call to TABULATEPOPULATIONGROUP by $s = 9$.

We summarize the parameters for the two versions of the SafeTab alongside the MOE targets in in Table 2. In this table, we also calculate the total privacy loss of step 2 tabulation over all population groups at a population group level. This loss is 90% of the total loss.

Note that while the inputs to the SafeTab[Geometric] and SafeTab[Discrete Gaussian] are the pure dp and zCDP privacy losses respectively, it is also possible to analyze the algorithms under different privacy guarantees (e.g. in the next section we analyze the geometric version of the algorithm using RDP).

8.2 Privacy loss comparison approach

In this section we describe multiple analyses of the privacy loss of SafeTab[Discrete Gaussian] and SafeTab[Geometric]. Results appear Table 3. For each analysis except the pure differential privacy analysis (where $\delta = 0$), we use the approximate differential privacy loss with $\delta = 10^{-10}$.

Pure DP loss of SafeTab[Geometric] To compute the pure differential privacy loss for SafeTab[Geometric], we can simply take the sum of $\rho_1, \ldots \rho_7$ by the analysis in Section 6.1.

Approximate DP loss of SafeTab[Geometric] using RDP analysis From the analysis in Section 6.2, we know that the SafeTab[Geometric] algorithm satisfies $(\alpha, f(\alpha))$-RDP, where

$$f(\alpha) = \sum_{i=1}^{\omega} \left[s \cdot \max \left[\tau\left(\alpha, \frac{\gamma\rho_i}{s}\right) + \tau\left(\alpha, \frac{(1-\gamma)\rho_i}{s}\right), \tau\left(\alpha, \frac{\rho_i}{s}\right) \right] \right] \tag{39}$$

We next convert this guarantee to an approximate DP guarantee. Using Lemma 4, we have that

$$\epsilon = f(\alpha) + \frac{\log(1/\delta) + (\alpha - 1)\log(1 - 1/\alpha) - \log(\alpha)}{\alpha - 1}. \tag{40}$$

Setting $\delta = 10^{-10}$, we can find the optimal value of ϵ by minimizing the right hand expression over $\alpha \in (0, \infty)$. This minimum is challenging to find analytically we compute the expression for $\alpha \in \{1.01, 1.02, \ldots, 9.99, 10.0\}$ and take the minimum.

Approximate DP loss of SafeTab[Discrete Gaussian] using zCDP experimental analysis From the analysis in Section 7.1, SafeTab[Discrete Gaussian] satisfies ρ-zCDP where $\rho = \sum_{i=1}^{7} \rho_i$. We can convert this into an approximate DP guarantee using equation 6. Rather than computing the infimum analytically, we compute the expression for $\alpha \in \{1.01, 1.02, \ldots, 9.99, 10.0\}$ and take the minimum.

Base Mechanism	Geometric		Discrete Gaussian	
Analysis	Pure DP	RDP	zCDP (analytical)	zCDP (experimental)
$(\epsilon, 10^{-10})$ **privacy loss**	15.3 ($\delta = 0$)	13.2	12.8	12.2

Table 3: The approximate differential privacy loss of the SafeTab[Geometric] and SafeTab[Discrete Gaussian] algorithm for various analyses. For all but the Pure DP analysis, $\delta = 10^{-10}$. For the Pure DP analysis, $\delta = 0$.

Approximate DP loss of SafeTab[Discrete Gaussian] using zCDP analytic analysis In addition to the experimental analysis, we also convert the ρ-zCDP to approximate DP analytically using Lemma 5.

8.3 Results & Discussion

Privacy losses of Geometric vs Discrete Gaussian: Results appear in Table 3. We observe the following key findings:

- The pure DP privacy loss of SafeTab[Geometric] is bounded by $\epsilon = 15.3$. This is smaller than the $\epsilon = 18$ that was analyzed by the Census POP team using the SafeTEx analysis tool. The lower privacy loss is due to two reasons:
  1. The original SafeTEx analysis was done with a target MOE of 5.6 for Nation and State detailed population groups. We also updated the MOE to 6 since noise is integral and therefore it makes sense to use an integral MOE.
  2. The original SafeTEx analysis was done assuming SafeTab used the base Laplace mechanism rather than the base geometric mechanism. The tail probability of the Geometric distribution are tighter than those of the Laplace distribution at integer points. This results in a smaller privacy loss.
- As expected SafeTab[Geometric] permits a smaller privacy loss (ϵ) under approximate DP with ($\delta = 10^{-10}$). We are able to achieve this by analyzing SafeTab[Geometric] under Rényi DP.
- We observe that the privacy loss of SafeTab[Discrete Gaussian] is smaller than that of SafeTab[Geometric] (when $\delta = 10^{-10}$). The improvement in privacy loss is small (3% for the analytical bound and 7% for the experimental bound).

Granularity of Statistics under Geometric vs Discrete Gaussian: The granularity of statistics released by SafeTab depends on the thresholds used as well as the noise scale used in Step 1 of the adaptive part of the algorithm. Under the MOE settings in Table 2 the noise scales for Step 1 should be roughly the same under both SafeTab[Geometric] and SafeTab[Discrete Gaussian]. So we should expect statistics to be released at roughly the same granularity. A quantitative analysis requires running these algorithms on either the simulated or 2010 data, which we defer to future work.

MOE for (Nation, Detailed) and (State, Detailed)	Geometric		Discrete Gaussian	
	Pure DP	RDP	zCDP (analytical)	zCDP (experimental)
5	16.7 ($\delta = 0$)	14.6	15.0	14.3
6	15.3 ($\delta = 0$)	13.2	12.8	12.2
7	14.2 ($\delta = 0$)	12.1	11.3	10.7
8	13.4 ($\delta = 0$)	11.3	10.2	9.7
9	12.7 ($\delta = 0$)	10.7	9.5	9.0
10	12.2 ($\delta = 0$)	10.2	8.9	8.4
11	11.7 ($\delta = 0$)	9.7	8.4	8.0

Table 4: The approximate differential privacy loss of the SafeTab[Geometric] and SafeTab[Discrete Gaussian] algorithm when $\delta = 10^{-10}$ for alternate MOE values for the (Nation, Detailed) and (State, Detailed) population group levels. MOEs for all other population group levels are fixed as in Table 2.

Alternate MOEs for Nation and State Detailed counts: It is evident from Table 2 that most of the privacy loss results from releasing detailed counts for Nation and State population groups. We present in Table 4 the overall privacy loss that might result from changing the MOE slightly for (Nation, detailed) and (State, detailed) population group levels from 6 to 5, 7, 8, 9, 10 and 11. We did not increase beyond 11, as that is the MOE for County and AIANNH detailed population groups. We note that increasing the MOE has a significant impact on the privacy loss. The zCDP (experimental) privacy loss bounds drops from around 12 to around 8, about a third reduction in the privacy loss.

References

[1] U.s. code title 13—census. https://www.law.cornell.edu/uscode/text/13.

[2] Memorandum 2019.25: 2010 demonstration data products – design parameters and global privacy-loss budget. https://www.census.gov/programs-surveys/decennial-census/2020-census/planning-management/memo-series/2020-memo-2019_25.html, October 2019.

[3] Seth Borenstein. Potential privacy lapse found in americans' 2010 census data. https://apnews.com/article/aba8e57c145047b5bab11b62baaa7f7a, February 2019.

[4] Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. *CoRR*, abs/1605.02065, 2016.

[5] Clément L. Canonne, Gautam Kamath, and Thomas Steinke. The discrete gaussian for differential privacy. *CoRR*, abs/2004.00010, 2020.

[6] Frank McSherry. Privacy integrated queries: an extensible platform for privacy-preserving data analysis. In Ugur Çetintemel, Stanley B. Zdonik, Donald Kossmann, and Nesime Tatbul, editors, *Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD 2009, Providence, Rhode Island, USA, June 29 - July 2, 2009*, pages 19–30. ACM, 2009.

[7] Ilya Mironov. Rényi differential privacy. In *30th IEEE Computer Security Foundations Symposium, CSF 2017, Santa Barbara, CA, USA, August 21-25, 2017*, pages 263–275, 2017.